Prisoners' Other Dilemma*

Matthias Blonski† Giancarlo Spagnolo‡

July 24, 2001

(SSE/EFI WP No. 437, *http://swopec.hhs.se*)

Abstract

We find that – contrary to common perception – cooperation as equilibrium of the infinitely repeated discounted Prisoner's Dilemma is in many relevant cases not very plausible: For a significant subset of the payoff-discount factor parameter space, *all* cooperation equilibria are strictly *risk dominated* (in the sense of Harsanyi and Selten, 1988) by non-cooperation. We derive an easy-to-calculate critical level for the discount factor δ^* below which this happens, and argue it is a better measure for the "likelihood" of cooperation than the critical level $\underline{\delta} < \delta^*$ at which cooperation is supportable in equilibrium. The results apply to other games sharing the strategic structure of the Prisoner's Dilemma (implicit/relational contracts, public goods games, etc.). We illustrate our main result for collusion equilibria in the repeated Cournot duopoly.

JEL Classification Numbers: C72

*We appreciate suggestions and remarks by Eric van Damme, Christian Ewerhart, David Levine, Peyton Young and various seminar participants in Berkeley, Mannheim, Stanford and Stockholm. This project has been supported by grants from the DFG (BL418/4-1) for Blonski, and from the European Commission (EGEMTPS/RTN) for Spagnolo.

†Department of Economics, University of Mannheim, Seminargebäude A 5, D-68131 Mannheim, Germany; telephone: +49-(0)621-181-1876; fax: +49-(0)621-181-1874. e-mail correspondence: *blonski@rumms.uni-mannheim.de*

‡University of Mannheim, Stockholm School of Economics, and CEPR.

1 Introduction

Consider the following specification of a Prisoner's Dilemma game, denoted by Γ, where $a < \frac{7}{3}$.

Γ	c	d
c	3, 3	a, $\frac{10}{3}$
d	$\frac{10}{3}$, a	$\frac{7}{3}$, $\frac{7}{3}$

Most students in social sciences are shown a similar game at some point in their first years, and are told the famous story about two prisoners invented by Albert Tucker in 1950 for a seminar in the psychology department in Stanford. They learn that Γ is a dominance solvable game and that the strategy profile $D := (d, d)$ is the unique pure strategy equilibrium point, although being strictly Pareto-dominated by the strategy profile $C := (c, c)$.

Later on in their curriculum, some of these students also learn that if players are sufficiently patient, this dilemma can be overcome by repeated interaction. The most frequently used model to formulate this way of "fixing the dilemma" is the infinitely repeated supergame with common discount factor δ, denoted by $\Gamma(\delta)$ (e.g. Friedman, 1971), where cooperation can be supported in subgame perfect equilibrium as long as players' discount factor is above the lower bound $\underline{\delta}$ that equalizes short run gains and (maximal) long run losses from defecting. In our example, cooperation is supportable as equilibrium behavior in $\Gamma(\delta)$ as long as $\delta \geq \underline{\delta} := \frac{1}{3}$.[1] In applications this lower bound $\underline{\delta}$ is often used as an (inverse) index of how plausible cooperation is in a given environment.

The literature on renegotiation-proofness has shown that this conclusion does not change if we require continuation strategies not to be Pareto-dominated (e.g. van Damme, 1989). What can one say more about such a well-understood game?

Our work originates from two simple observations: (i) The lower bound $\underline{\delta}$ does not depend on the parameter a of the stage game. (ii) The parameter a may or should influence players' propensity to cooperate. In this article we formalize the consequences of variations in the parameter a in the infinitely repeated discounted Prisoner's Dilemma, and in games with analogous features.

[1] Then discounted payoffs from playing cooperation indefinitely, $\frac{3}{1-\delta}$, offset those from defecting unilaterally and being kept at the minimax thereafter, $\frac{10}{3} + \frac{7}{3}\frac{\delta}{1-\delta}$.

To convince the reader that stage game parameter a should be taken into account, and to give a hint on how this can be done in a precise way, let us look at $\Gamma(\delta)$ for $a = -\frac{14}{3}$ and $\delta = \frac{2}{3} = 2\underline{\delta}$. Suppose for the moment that players only consider the following two pure strategies of the repeated stage game:

- c^* : Cooperate as long as no player defects, defect forever otherwise.
- d^* : Defect forever.

The so defined $2 \times 2-$ game Γ^* is given by

Γ^*	c^*	d^*
c^*	9 9	0 8
d^*	8 0	7 7

,

where numbers are discounted sums of payoffs, and reflect the strategic reasoning of players playing $\Gamma(\delta)$ but restricting attention only to the two strategies c^* and d^*. The game Γ^* has two strong pure strategy equilibria $C^* := (c^*, c^*)$ and $D^* = (d^*, d^*)$ where D^* is strictly payoff dominated by C^*.[2] Which of those two equilibria do we expect to be selected?

While $C^* := (c^*, c^*)$ is the payoff-dominant equilibrium, a cautious player might prefer to play d^* to avoid extreme losses, for example if the opponent makes mistakes. But even rational (and never failing) players who just are not sure about their opponent's beliefs on their own rationality, or have any higher order doubt, may prefer d^* over c^*. Moreover, pre-play communication is of very limited value in this game since a player planning to play d^* has an incentive to convince his opponent to play c^*.

Harsanyi and Selten, in their monumental book on equilibrium selection (1988, subsequently abbreviated with HS) introduce the selection criterion *risk dominance* and demonstrate that in $2 \times 2-$games with two strong pure strategy equilibrium points the risk dominance relation (in contrast to the payoff dominance relation) is invariant under

[2]There is another (weak) mixed equilibrium which is not of interest here.

transformations on payoffs preserving the best reply structure.[3] The game Γ^{**} given by

Γ^{**}	c^{**}	d^{**}
c^{**}	1, 1	0, 0
d^{**}	0, 0	7, 7

is equivalent to Γ^* with respect to the best reply structure. Risk dominance in simple unanimity games is evaluated by comparing the so called Nash-products of the two equilibria. We picked these numbers for the example because Γ^* has made a career into game theory textbooks (see for example Fudenberg and Tirole, p. 21) as an example of a game where the selection criteria payoff dominance and risk dominance conflict with each other.

The reduced game Γ^* only captures a tiny part of the huge supergame with its universe of equilibria. In this article we investigate for which PD games this "risk-dominance problem" is present for *all* equilibria supporting cooperation, and lable this *Prisoners' Other Dilemma*. We characterize the set of PD-games parameters featuring Prisoners' Other Dilemma, and show that it is never empty (in particular, for no given discount factor). More importantly for economists, this set includes relevant cases as the textbook example of a collusion game for Cournot-duopolists facing a linear demand function.

We believe that when risk dominance and payoff dominance conflict with each other $\underline{\delta}$ is no longer a good indicator for players' inclination towards cooperative behavior as it is used in the applied literature. As one of our results, we propose an alternative lower bound for the discount factors δ^*, with $\delta^* > \underline{\delta}$, that is a function of parameter a and reflects the riskiness to cooperate besides the incentive to defect (in our introductory example, $\delta^* = \frac{11}{12} > \frac{1}{3}$).

Finally, we argue that if players are susceptible to risk considerations, an equilibrium should be considered safe only if it is not risk dominated in any of its (out-of-equilibrium) subgames. We name this property *risk perfection*, and provide sufficient conditions under which it is satisfied for cooperation equilibria. It turns out that these conditions include the relevant punishment strategies being used in the (theoretical and applied) literature.

[3] Payoff differences (incentives to switch to another strategy) for any given opponent's behavior remains unchanged.

Although HS favour payoff-dominance over risk dominance as selection criterion, the theoretical and experimental support for risk dominance has increased since then. Theoretical support has been offered by the evolutionary game theory literature (see for example Kandori, Mailath and Rob, 1993; and Young 1993) and by the literature on global games considering perturbations of payoff parameters (starting with Carlsson and van Damme, 1993). Experimental evidence also strongly supports risk and security if they conflict with the payoff criterion (see for example van Huyck, Battalio, and Beil, 1990).

The HS-concept of risk-dominance as pairwise comparison between two equilibrium points is difficult for general games and did not find its way into the applied literature. HS axiomatized it only for the rather limited class of $2\times2-$games with two strict equilibrium points (see HS, Ch. 3). In this article we demonstrate that for other economically meaningful cases risk dominance is easy to define and yields immediate insights which, we believe, are difficult to obtain otherwise.

We focus in this paper on the discounted infinitely repeated Prisoner's Dilemma. However, definitions and results directly apply to many other games that share its strategic features, including repeated oligopoly and public good games, and implicit/relational contracting models.

2 Risk dominance

In this section we consider the following symmetric[4] PD stage game Γ characterized by payoff parameters[5] a, b, c, d where $b > c > d > a$ and $2c > b + a$,

Γ	c	d
c	c, c	a, b
d	b, a	d, d

.

Denote the vector of payoff parameters by $\lambda = (a, b, c, d)$ and together with a common discount factor by $s = (\lambda, \delta) = (a, b, c, d, \delta)$ and call $\Gamma(s)$ the related infinitely repeated

[4]All definitions and results are very similar if the symmetry assumption is dropped. We use this assumption mainly for notational convenience (four payoff parameters instead of eight).

[5]In order to economize on notation we use c, d as lables for the stage game strategies as well as for payoff parameters as long as there is no confusion.

game with common discount factor δ. We call an equilibrium φ that supports indefinite cooperation as its equilibrium outcome path a *cooperation equilibrium.* Denote by ω the max-min equilibrium or *defection equilibrium* where each player plays the unique stage game dominant strategy forever.

The prime question of the present inquiry is: Under which circumstances – i.e. for which parameter constellations s – are cooperation equilibria not very plausible?[6] Since in this article we adopt risk dominance as selection criterion, our first task will be to define risk dominance in a simple way within this specific class of repeated games. HS define risk dominance for a very general class of games (extensive games with perfect recall in standard form) as a pairwise comparison between two strong equilibrium points.[7] Here, we proceed by introducing a simple and intuitive definition of the risk dominance criterion applied to the pair φ, ω of equilibria for any cooperative equilibrium φ. This definition is specifically adjusted to the problem under investigation and to the very nature of the game $\Gamma(\delta)$. We then show that our definition is equivalent to the more general HS-definition.

For this purpose we introduce, for each cooperation equilibrium φ, the related but much simpler game Γ_φ – subsequently called the $\varphi-$*formation of* $\Gamma(\delta)$. The $\varphi-$formation Γ_φ is a substructure of $\Gamma(\delta)$ capturing the strategic considerations of players restricting their attention to the binary subset of the strategy space defined by the two equilibria φ and ω. More precisely, Γ_φ is the $2\times 2-$game defined by the strategy space $\{\varphi_i, \omega\}$ where $\varphi = (\varphi_1, \varphi_2)$ and $\omega = (\omega, \omega)$ are the equilibrium strategy profiles for both players.[8] The bimatrix-form of Γ_φ is given by

Γ_φ	φ_2	ω
φ_1	$\frac{c}{1-\delta}$, $\frac{c}{1-\delta}$	$a+\delta V_{1\varphi}$, $b+\delta V_{2\omega}$
ω	$b+\delta V_{1\omega}$, $a+\delta V_{2\varphi}$	$\frac{d}{1-\delta}$, $\frac{d}{1-\delta}$

,

where $V_{i\xi}$ for $i\xi \in \{1\varphi, 1\omega, 2\varphi, 2\omega\}$ are the equilibrium payoffs of the corresponding

[6]Obviously the same question can be asked for every other equilibrium of the repeated PD game, even for inefficient equilibria. We will demonstrate in the extensions that the essence of Prisoners' Other Dilemma remains the same or even aggravates by moving away from full cooperation.

[7]In HS an equilibrium point is strong if every player's equilibrium strategy is a strong best reply.

[8]In the defect equilibrium both players use the same strategy. As long as this does not cause confusion we identify the strategy and the corresponding equilibrium profile with the same symbol ω.

continuation games.[9]

Denote by $\psi = (\psi, \psi)$ the 'grim-trigger'- or 'Nash reversion'-punishment cooperation equilibrium where every player responds to a deviation from cooperation by defecting forever.

For $\delta < \underline{\delta} := \frac{b-c}{b-d}$ the cooperation equilibrium set is empty since

$$b + \delta V_{i\omega} \geq b + \delta V_{\psi} = b + \frac{\delta d}{1-\delta} > \frac{c}{1-\delta},$$

and even the most severe punishment cannot support indefinite cooperation as equilibrium point of $\Gamma(s)$. To rule out these uninteresting cases and cases where cooperation can only be supported by weak equilibria let us assume from now $\delta > \underline{\delta}$ and denote the respective parameter space by

$$S := \{s = (a, b, c, d, \delta) \,| b > c > d > a, \, 2c > b + a \text{ and } \underline{\delta} < \delta < 1\}.$$

Definition 1 *Let Φ denote the set of cooperation equilibria such that the $\varphi-$formation game Γ_{φ} has two strong pure-strategy equilibrium points, i.e.*

$$\begin{aligned} u_i(\varphi) \; &: = \frac{c}{1-\delta} - b - \delta V_{i\omega} > 0 \quad \textit{and} \\ v_i(\varphi) \; &: = \frac{d}{1-\delta} - a - \delta V_{i\varphi} > 0, \end{aligned}$$

for $i = 1, 2$.

Lemma 1 *Let φ be a cooperation equilibrium that is not in Φ. Then φ is a weak equilibrium. In particular, some player i is indifferent between his equilibrium strategy φ_i and defecting forever ω.*

Proof. By definition of Φ we get $\varphi \notin \Phi \Rightarrow u_i(\varphi) \leq 0$ or $v_i(\varphi) \leq 0$ for some i. In order to be an equilibrium φ must satisfy $u_i(\varphi) \geq 0$ for $i = 1, 2$. Further, $V_{i\varphi} \leq \frac{d}{1-\delta} < \frac{1}{\delta}\left(\frac{d}{1-\delta} - a\right)$ implies $v_i(\varphi) > 0$ for $i = 1, 2$. This together with $\varphi \notin \Phi$ yields $u_i(\varphi) = 0 \Leftrightarrow \frac{c}{1-\delta} = b + \delta V_{i\omega}$ for some i. ■

¿From here we restrict attention to the set Φ of strong cooperation equilibria since weak equilibria are even more "risky" in the sense to be defined now. To see that Φ is always non-empty note that the corresponding maximal payoff-differences $u(\psi), v(\psi)$

[9] For the present purpose all cooperation strategies φ defining the same vector $V = (V_{1\varphi}, V_{1\omega}, V_{1\varphi}, V_{1\omega})$ can be identified since all further definitions and results do not depend on deviations further down the game tree.

for the grim-trigger-punishment do not depend on i and for our parameter restrictions are always strictly positive since

$$\begin{aligned} u(\psi) &= \frac{c}{1-\delta} - b - \frac{d\delta}{1-\delta} = \frac{\delta(b-d)-(b-c)}{1-\delta} > 0 \quad \text{and} \\ v(\psi) &= \frac{d}{1-\delta} - a - \frac{d\delta}{1-\delta} = d - a > 0. \end{aligned}$$

Every $2 \times 2-$game with two strong equilibrium points is equivalent with respect to the best response structure to a simple unanimity game.[10] We denote by Γ^*_φ the unanimity game being best-response-equivalent to Γ_φ. The bi-matrix form of Γ^*_φ is given by

Γ^*_φ	φ^*_2	ω^*
φ^*_1	$u_1(\varphi)$, $u_2(\varphi)$	0, 0
ω^*	0, 0	$v_1(\varphi)$, $v_2(\varphi)$

.

The formulation *unanimity-game* suggests that in any equilibrium both players must agree to choose the 'same' strategy. Now we are ready to define risk dominance.

Definition 2 *We call a cooperation equilibrium $\varphi \in \Phi$ "strictly risk dominated" by the defection equilibrium ω iff*

$$u_1(\varphi)\, u_2(\varphi) < v_1(\varphi)\, v_2(\varphi),$$

and write $\omega \succ_{rd} \varphi$. The corresponding weak relation is denoted by $\succsim_{rd}$.

Harsanyi and Selten sometimes call $u_1(\varphi)\, u_2(\varphi)$ and $v_1(\varphi)\, v_2(\varphi)$ 'Nash-products' of the corresponding equilibria. Hence, we can say, risk dominance relies on comparing Nash-products of the equilibria in the related formation.

Lemma 2 *Risk dominance as defined here is equivalent to the more general definition of risk dominance applied to repeated PD-games of the type $\Gamma(s)$ as given in Harsanyi and Selten (1988).*

Proof. See Appendix. ■

[10]Adding numbers to player 1's payoff columns and to player 2's payoff rows does not alter both players' deviation incentives and – in the language of HS – preserves the best response structure.

Definition 3 *We call $\rho(s,\varphi)$ the "riskiness" of cooperation equilibrium φ in $\Gamma(s)$, where*

$$\begin{aligned}\rho(s,\varphi) &: = v_1(\varphi)\, v_2(\varphi) - u_1(\varphi)\, u_2(\varphi) \\ &= \left(\frac{d}{1-\delta} - a - \delta V_{1\varphi}\right)\left(\frac{d}{1-\delta} - a - \delta V_{2\varphi}\right) \\ &\quad - \left(\frac{c}{1-\delta} - b - \delta V_{1\omega}\right)\left(\frac{c}{1-\delta} - b - \delta V_{2\omega}\right).\end{aligned}$$

The condition $\rho(s,\varphi) > 0$ exactly characterizes all strictly risk dominated cooperation equilibria. For any given cooperation equilibrium φ of a discounted PD-supergame $\Gamma(s)$ definition 3 can be easily applied to verify whether the equilibrium is risk dominated. The same can be done for cooperation equilibria of other discounted supergames with analogous strategic features.

3 Characterization of Prisoners' Other Dilemma

Since $c > d$, any cooperation equilibrium φ payoff-dominates defection ω. Our previous definition shows that risk-dominance may point to the opposite direction for a particular cooperation equilibrium. Our task in this section is to characterize the set of all strong cooperation equilibria where payoff dominance and risk dominance point to opposite directions and vice versa. We begin by establishing an important benchmark.

Proposition 1 *There is no cooperation equilibrium $\varphi \in \Phi$ which is less risky than the grim trigger equilibrium ψ. Formally,*

$$\rho(s,\psi) = \underline{\rho}(s) := \inf_{\varphi} \rho(s,\varphi).$$

Proof. No cooperation equilibrium φ can be less risky than $\inf_{\varphi} \rho(s,\varphi)$. First, note that for any $\varphi \in \Phi$ the upper bound for the continuation payoff of a player who plays a cooperation equilibrium strategy against a player who always defects is

$$V_{i\varphi} \leq \frac{d}{1-\delta},$$

and that continuation payoffs of players who always defect are bounded by

$$\frac{d}{1-\delta} \leq V_{i\omega} \leq \frac{b}{1-\delta}.$$

Next, to consider strong cooperation equilibria in Φ imposes additional boundaries on $V_{i\varphi}$ and $V_{i\omega}$ given by $V_{i\varphi} < \frac{1}{\delta}\left(\frac{d}{1-\delta} - a\right)$ and $V_{i\omega} < \frac{1}{\delta}\left(\frac{c}{1-\delta} - b\right)$. Only the second inequality is binding, hence together with the boundaries for $\Gamma(\delta)$ we obtain

$$\begin{aligned} V_{i\varphi} &\leq \frac{d}{1-\delta} \text{ and} \\ \frac{d}{1-\delta} &\leq V_{i\omega} < \frac{1}{\delta}\left(\frac{c}{1-\delta} - b\right), \end{aligned}$$

since $b > c > d > a$ implies $\frac{d}{1-\delta} < \frac{1}{\delta}\left(\frac{d}{1-\delta} - a\right)$ and $\frac{b}{1-\delta} > \frac{1}{\delta}\left(\frac{c}{1-\delta} - b\right)$. Then continuation payoffs must satisfy $V_{i\varphi} \leq \frac{d}{1-\delta}$ and $V_{i\omega} \geq \frac{d}{1-\delta}$. For the grim-trigger strategy equilibrium ψ both conditions are binding and hold with equality. This yields

$$\begin{aligned} \underline{\rho}(s) &= \inf_{\varphi} \rho(s,\varphi) \\ &= \left(\frac{d}{1-\delta} - a - \delta\frac{d}{1-\delta}\right)^2 - \left(\frac{c}{1-\delta} - b - \delta\frac{d}{1-\delta}\right)^2 \\ &= (d-a)^2 - \left(\frac{c}{1-\delta} - b - \delta\frac{d}{1-\delta}\right)^2 \\ &= \rho(s,\psi). \end{aligned}$$

■

It is now time to state our theorem, the main result of this paper. In order to be more precise about parameters, we introduce the following notation.

Definition 4 *Let S^ω denote the set of repeated PD-games where all strong cooperation equilibria are strictly risk dominated by the defection equilibrium, i.e.*

$$S^\omega := \{s \in S \,|\, \omega \succ_{rd} \varphi \ \forall \varphi \in \Phi(s)\} \subset S.$$

Conversely, let S^φ denote the set of repeated PD-games where no strong cooperation equilibrium is strictly risk dominated by the defection equilibrium, i.e.

$$S^\varphi := \{s \in S \,|\, \varphi \succsim_{rd} \omega \ \forall \varphi \in \Phi(s)\} \subset S.$$

The following theorem characterizes these parameter sets.

Theorem 1 *(i) For $\delta \in (\underline{\delta}, \delta^*)$ with $\delta^* := \frac{b-a-(c-d)}{b-a} > \underline{\delta}$ all cooperation equilibria of the repeated PD-game $\Gamma(s)$ are strictly risk dominated, hence $S^\omega = \{s \in S \,|\, \delta < \delta^*\}$.*

(ii) Conversely, there exist no parameters $s \in S$ such that no cooperation equilibrium is strictly risk dominated, hence $S^\varphi = \emptyset$.

Proof. A little calculation shows that the interval $(\underline{\delta}, \delta^*)$ is never empty

$$\begin{aligned} (d-a)(c-d) &> 0 \Leftrightarrow \\ \frac{b-a-c+d}{b-a} &> \frac{b-c}{b-d} \Leftrightarrow \\ \delta^* &> \underline{\delta}. \end{aligned}$$

This implies that for $\delta \in (\underline{\delta}, \delta^*)$

$$\begin{aligned} \delta &< \frac{b-a-c+d}{b-a} \Leftrightarrow \\ d-a &> \frac{c}{1-\delta} - b - \delta \frac{d}{1-\delta} \Leftrightarrow \\ \left(\frac{d}{1-\delta} - a - \delta \frac{d}{1-\delta}\right)^2 &> \left(\frac{c}{1-\delta} - b - \delta \frac{d}{1-\delta}\right)^2 \Leftrightarrow \\ \rho(s, \psi) &= \underline{\rho}(s) > 0. \end{aligned}$$

Since all implications hold in both directions this implies claim (i) of the theorem. To prove claim (ii), define similarly as in lemma (1)

$$\bar{\rho}(s) := \sup_{\varphi} \rho(s, \varphi)$$

as the lowest upper bound on the riskiness among all cooperation equilibria. It remains to show that $\bar{\rho}(s)$ is strictly positive $\forall s \in S$. By the boundaries given in the proof of proposition 1 we know that

$$\inf_{\varphi} \left(\frac{c}{1-\delta} - b - \delta V_{1\omega}\right)\left(\frac{c}{1-\delta} - b - \delta V_{2\omega}\right) = 0$$

and that

$$\sup_{\varphi} \left(\frac{d}{1-\delta} - a - \delta V_{1\varphi}\right)\left(\frac{d}{1-\delta} - a - \delta V_{2\varphi}\right) \in \left[(d-a)^2, \left(\frac{d-a}{1-\delta}\right)^2\right].$$

This together yields

$$\begin{aligned} \bar{\rho}(s) &= \sup_{\varphi} \left(\frac{d}{1-\delta} - a - \delta V_{1\varphi}\right)\left(\frac{d}{1-\delta} - a - \delta V_{2\varphi}\right) \\ &\quad - \inf_{\varphi} \left(\frac{c}{1-\delta} - b - \delta V_{1\omega}\right)\left(\frac{c}{1-\delta} - b - \delta V_{2\omega}\right) \\ &\geq (d-a)^2 - 0 \\ &> 0. \end{aligned}$$

■

We label *Prisoners' Other Dilemma* the problem that incentive compatible cooperative behavior ($\delta > \underline{\delta}$) may be considered too risky to "fix Prisoner's (original) Dilemma" in repeated interactions. The theorem tells us exactly when prisoners susceptible to risk dominance are unable to overcome the original dilemma by building up a "cooperative relationship". This is the case when $\delta < \delta^*$. The theorem moreover tells us that there exists no discounted repeated PD-game for which this "other dilemma" disappears altogether. There are always some risky cooperation equilibria. Intuitively, one obtains the more risky cooperation equilibria by letting players be "forgiving", i.e. try to start cooperative behavior although the opponent defected in the past. In equilibrium, however, this cannot be done too frequently.

The following corollary follows immediately from the theorem. It points to stage game parameter constellations where Prisoners' Other Dilemma tends to be most serious.

Corollary 1 *For a very large payoff-difference $b - a$ or a very small difference $c - d$ all cooperation equilibria are risk-dominated for almost any discount factor $\delta < 1$. Formally,*

$$\lim_{b-a \to \infty} \delta^* = \lim_{c-d \to 0} \delta^* = 1.$$

The following proposition has the flavour of an 'anti-Folk theorem for risk-dominance' and shows that for any discount factor $\delta < 1$ and appropriately chosen payoff parameters all cooperation equilibria are strictly risk dominated. Moreover by choosing the payoff parameter a sufficiently low the riskiness of all cooperation equilibria can be made arbitrarily large.

Proposition 2 *For every $\delta < 1$ there exist payoff parameters λ with $s = (\lambda, \delta) \in S^\omega$ such that all cooperation equilibria of $\Gamma(s)$ are strictly risk dominated. Moreover, for any given riskiness $\rho > 0$ there exist payoff parameters λ with $s = (\lambda, \delta) \in S^\omega$ such that all cooperation equilibria have at least riskiness ρ.*

Proof. Payoff parameter a is not bounded from below. Hence

$$\underline{\rho}(s) = (d-a)^2 - \left(\frac{c}{1-\delta} - b - \delta\frac{d}{1-\delta}\right)^2$$

goes to infinity for $a \to -\infty$. This implies both statements of the proposition. ■

4 Risk perfection

The idea that in a repeated Prisoner's Dilemma game $\Gamma(s)$ players might consider a cooperation equilibrium 'too risky' – although it Pareto-dominates other equilibria – carries over in a natural way to the subgames of $\Gamma(s)$. If players are susceptible to risk dominance, they are so at all nodes of the game, hence a risk undominated equilibrium path supported by risk dominated out-of-equilibrium (punishment) paths may not be considered a 'safe' equilibrium. The occurence of past deviations might even reinforce risk considerations, further favoring risk dominance as selection criterium.

A subgame $\Gamma^h(s)$ of $\Gamma(s)$ is characterized by a history $h \in H$ specifying the path of stage game actions up to the period where the subgame starts. Risk-dominance of a cooperation equilibrium φ^h and riskiness $\rho^h(s,\varphi)$ restricted to $\Gamma^h(s)$ are defined equivalently by comparing Nash-products in the corresponding formation $\Gamma^h_\varphi(s)$, hence we can introduce the following refinement.

Definition 5 *A cooperation equilibrium $\varphi \in \Phi(s)$ is called* risk perfect *iff its restriction to any subgame is not strictly risk dominated. Formally:*

$$\rho^h(s,\varphi) \leq 0 \quad \forall h \in H$$

It is easy to recognize that the grim trigger equilibrium ψ is risk perfect whenever it is not strictly risk dominated. After any deviation ψ the stage game equilibrium is played forever, which is perfectly safe at any later instant. Hence, the condition $\delta \geq \delta^*$ also guarantees that at least one risk perfect cooperation equilibrium exists.

Which other equilibria are risk perfect? To give sufficient conditions for risk perfection we restrict attention to *simple strategies* as defined by Abreu (1988). In the 2-player repeated Prisoner's Dilemma a simple strategy for player i is specified by 3 paths, the initial path π^0 and a punishment path π^j for every player $j = 1, 2$. A punishment path specifies what is played if player j deviates from the initial path or any ongoing punishment path. If no player deviates or both players deviate simultaneously a simple strategy specifies to proceed along the ongoing path.[11] As Abreu showed, every perfect equilibrium outcome can be supported by a perfect equilibrium in simple strategies.

[11] To avoid introducing further notation we do not provide a formal definition of simple strategies and optimal penal codes. The theory is well known, and we do not need it here; for more details, please, see Abreu (1988).

Definition 6 *We call a punishment path π^j of a simple strategy in the repeated Prisoner's Dilemma a* monotonous restitution *if (i) no player ever switches from c to d along the path (monotony) and (ii) the punishing party $i \neq j$ never switches from d to c before the reneging party j does (restitution).*

A monotonous restitution after a deviation of, say, player 1 always takes the form

$$\pi^1 = \left(\underbrace{(d,d), ..., (d,d)}_{\text{punishment phase: } T_1 \text{ periods}}, \underbrace{(c,d), ..., (c,d)}_{\text{restitution phase: } \tau_1 \text{ periods}}, \underbrace{(c,c), (c,c), ...}_{\text{cooperation phase}} \right).$$

In a monotonous punishment path a player who starts to cooperate will cooperate forever. For example, the path $\pi^1 = ((d,d), (c,d), (d,d), (c,c), (c,c), ...)$ is clearly not monotonous, and $\pi^1 = ((d,d), (d,c), (c,c), ...)$ is not a restitution since the deviating player 1 starts cooperating later, gaining again instead of (weakly) recompensating his opponent. Monotonous restitutions include most punishments used in applications, among which:

- Grim trigger $T_i = \infty$:

$$\pi^1 = \pi^2 = ((d,d), (d,d), ...)$$

- Tit for tat, $T_i = 0, \tau_i = 1$:

$$\begin{aligned} \pi^1 &= ((c,d), (c,c), (c,c), ...) \\ \pi^2 &= ((d,c), (c,c), (c,c), ...) \end{aligned}$$

- T-periods "defection wars" or 0-restitution, $T > 0, \tau_i = 0$:

$$\pi^1 = \pi^2 = \left(\underbrace{(d,d), ..., (d,d)}_{T \text{ periods}}, (c,c), (c,c), ... \right)$$

- van Damme's (1989) renegotiation proof strategies, $T = 0, \tau_i > 0$:

$$\begin{aligned} \pi^1 &= \left(\underbrace{(c,d), ..., (c,d)}_{\tau \text{ periods}}, (c,c), (c,c), ... \right) \\ \pi^2 &= \left(\underbrace{(d,c), ..., (d,c)}_{\tau \text{ periods}}, (c,c), (c,c), ... \right). \end{aligned}$$

We can now state the following.

Proposition 3 *Consider the discounted infinitely repeated Prisoner's Dilemma* $\Gamma(s)$*. Let* φ *be a subgame perfect risk undominated cooperation equilibrium in simple strategies with monotonous restitution punishment paths. Then* φ *is risk perfect.*

Proof. To see that a subgame perfect risk undominated simple strategy equilibrium φ with monotonous restitution punishment paths is risk perfect we take advantage of the simple strategy concept. The subgame starting from any period in any future cooperation phase is equivalent to the initial path π^0. Hence assuming that on the initial path π^0 it is $\rho(s, \varphi) \leq 0$, we only need to verify that the same holds for all subgames $\Gamma_\varphi^h(s)$ starting within the monotonous restitution punishment paths π^i. To do this, we first identify some "critical" subgames, such that if $\rho(s, \varphi) \leq 0$ for that subgame, then $\rho(s, \varphi) \leq 0$ for all other subgames beginning in π^i. Then we verify risk dominance for the critical subgames.

Consider the monotonous restitution

$$\pi^1 = \left(\underbrace{(d,d), ..., (d,d)}_{\text{punishment phase: } T_1 \text{ periods}}, \underbrace{(c,d), ..., (c,d)}_{\text{restitution phase: } \tau_1 \text{ periods}}, \underbrace{(c,c), (c,c), ...}_{\text{cooperation phase}} \right)$$

where after $T_1 \geq 0$ periods of mutual non-cooperation the formerly defecting party "reimburses" the punishing party by unilaterally cooperating for τ_1 periods, with $\tau_1 \geq 0$. We called the first phase "punishment phase" and the second one "restitution phase". We now distinguish between the two cases (i) strict restitution: $\tau_1 \geq 1$ and (ii) 0-restitution: $\tau_1 = 0$.

Case (i) "Strict restitution": For $\tau_1 \geq 1$ the critical subgame starts at the beginning of the restitution phase in period $T_1 + 1$ of the monotonous restitution π^1. To see this, note that in subgames starting within the punishment phase sticking to equilibrium strategies is strictly less risky than in subgames starting during the restitution phase, since playing d involves no risk and players discount future (risk). Now note that the risk for player 1 involved in playing c at the beginning of the restitution phase $(T+1)$ is at least as large as in subsequent periods $(T_1 + 2)$ to $(T_1 + \tau_1)$. Therefore, for $\tau_1 \geq 1$ it remains to show that the risk dominance property is satisfied in the critical subgame beginning period $T_1 + 1$ of the monotonous restitution π^1, denoted by $\Gamma_\varphi^{h_1}(s)$. By our definition of risk dominance we have to look at the $2 \times 2-$formation $\Gamma_\varphi^{*h_1}(s)$ of $\Gamma_\varphi^{h_1}(s)$ where each player i only compares playing the equilibrium strategy $\varphi_i^{*h_1}$ and $\omega_i^{*h_1}$ (play always d). By subgame perfection this formation again must have two equilibria φ^{*h_1} and ω^{*h_1} induced by φ and ω. Next, note that a strict restitution phase $\tau_1 \geq 1$ prescribes that in φ^{*h_1} player 1 starts to reimburse player 2. If player 1, however, plays $\omega_1^{*h_1}$ and fails to do so, both players obtain the same payoff as in equilibrium ω^{*h_1} of the formation $\Gamma_\varphi^{*h_1}(s)$. Hence, ω^{*h_1} is a weak equilibrium since player 2 is indifferent between $\varphi_2^{*h_1}$ and $\omega_2^{*h_1}$ if player 1 plays $\omega_2^{*h_1}$. This implies that the Nash product of ω^{*h_1} is 0 and therefore is φ^{*h_1} not risk dominated.

Case (ii): "0-restitution phase": For the same reason as in the strict restitution phase the critical subgame is the one that starts at the beginning of the cooperation phase in the period T_1+1 of the monotonous restitution π^1. But this subgame is equivalent to the initial game starting in π^0 where the equilibrium φ is not risk dominated by assumption. This concludes the proof. ■

Hence, for most punishment strategies used in the literature (monotone restitutions), checking that the initial equilibrium path is not risk dominated ($\rho(s,\varphi) \leq 0$) is sufficient to guarantee risk perfection ($\rho^h(s,\varphi) \leq 0 \;\; \forall h \in H$). Regarding other equilibria, one has to check case by case. Consider, for example, an equilibrium in non-simple strategies where the first deviation from the equilibrium outcome path is punished differently than further deviations. Let φ be a cooperation equilibrium where punishment paths after the first deviation of player j, denoted by π_1^j, are given by

$$\pi_1^1 = \pi_1^2 = \left(\underbrace{(d,d),...,(d,d)}_{T^1 \text{ periods}}, (c,c), (c,c), ... \right),$$

with $T^1 > 1$. Now let $k(h)$ be the number of previous deviations from equilibrium behavior in history h, and suppose equilibrium strategies prescribe, for any further deviation $k > 1$,

$$\pi_k^1 = \pi_k^2 = \left(\underbrace{(d,d)}_{T^k=1}, (c,c), (c,c), ... \right),$$

i.e. defecting just once before returning to cooperation. Riskiness at the start of the game can be kept small by increasing T^1, while the subgame starting after these T^1 periods of punishment is subject to higher risk. It is easy to check that $T^1 > 1$ implies $\rho(s,\varphi) < \rho^k(s,\varphi)$ for $k > 1$. Hence, if parameters are such that $\rho(s,\varphi) \leq 0 < \rho^k(s,\varphi)$ the equilibrium is risk undominated but not risk perfect (it is risk perfect iff $\rho^k(s,\varphi) \leq 0$).

5 Extensions

5.1 Risk dominated efficient equilibria

One might wonder about the significance of Prisoners' Other Dilemma for efficient equilibria when payoffs are distributed asymmetrically among players on the Pareto frontier. Consider the regular case where symmetric cooperation equilibria are on the Pareto frontier ($2c > b + a$), and denote by $\theta(x)$ the efficient equilibrium yielding averaged asymmetric per-period payoffs $c - x, c + \frac{b-c}{c-a}x$ with $c - d \geq x \geq 0$. The following proposition is a reformulation of proposition 2 for these equilibria.

Proposition 4 *For every $\delta < 1$ there exist payoff parameters λ with $s = (\lambda, \delta) \in S$ such that all equilibria $\theta(x)$ supporting the same payoffs are strictly risk dominated. Moreover, for any given riskiness $\rho > 0$ there exist payoff parameters λ with $s = (\lambda, \delta) \in S$ such that all equilibria $\theta(x)$ have at least the riskiness ρ.*

Proof. The riskiness of $\theta = \theta(x)$ is given by

$$\begin{aligned} \rho(s,\theta) \quad : \quad &= v_1(\theta) v_2(\theta) - u_1(\theta) u_2(\theta) \\ &= \left(\frac{d}{1-\delta} - a - \delta V_{1\varphi}\right)\left(\frac{d}{1-\delta} - a - \delta V_{2\varphi}\right) \\ &\quad - \left(\frac{c-x}{1-\delta} - b - \delta V_{1\omega}\right)\left(\frac{c+\frac{b-c}{c-a}x}{1-\delta} - b - \delta V_{2\omega}\right) \end{aligned}$$

implying that $\rho(s,\theta) \to \infty$ for $a \to -\infty$. ■

The following proposition shows that moving away from symmetric cooperation along the Pareto frontier increases the riskiness if off-equilibrium punishments are kept constant and symmetric (as, for example, in 'grim trigger' and 'tit for tat').

Proposition 5 *Let φ be a cooperation equilibrium, and $\theta(x)$ be an equilibrium on the Pareto frontier yelding averaged asymmetric per-period payoffs $c - x, c + \frac{b-c}{c-a}x$, with $c - d \geq x > 0$, supported by the same off-equilibrium punishments as φ. Assume further that punishments are symmetric for defecting players $V_{2\omega} = V_{1\omega}$. Then $\theta(x)$ is more risky than φ and riskiness increases with x:*

$$\rho(s,\theta(x)) - \rho(s,\varphi) > 0 \text{ and } \frac{\partial}{\partial x}(\rho(s,\theta(x)) - \rho(s,\varphi)) > 0$$

Proof.

$$\begin{aligned}
\rho(s,\theta)-\rho(s,\varphi) &= v_1(\theta)\,v_2(\theta)-u_1(\theta)\,u_2(\theta)-(v_1(\varphi)\,v_2(\varphi)-u_1(\varphi)\,u_2(\varphi))\\
&= u_1(\varphi)\,u_2(\varphi)-u_1(\theta)\,u_2(\theta)\\
&= \left(\frac{c}{1-\delta}-b-\delta V_{1\omega}\right)\left(\frac{c}{1-\delta}-b-\delta V_{2\omega}\right)\\
&\quad -\left(\frac{c-x}{1-\delta}-b-\delta V_{1\omega}\right)\left(\frac{c+\frac{b-c}{c-a}x}{1-\delta}-b-\delta V_{2\omega}\right)\\
&= \left[\frac{c}{1-\delta}-b-\delta V_{2\omega}-\frac{b-c}{c-a}\left(\frac{c}{1-\delta}-b-\delta V_{1\omega}\right)\right]\frac{x}{1-\delta}\\
&= \left(\frac{c}{1-\delta}-b-\delta V_{1\omega}\right)\left(1-\frac{b-c}{c-a}\right)\frac{x}{1-\delta}\\
&> 0
\end{aligned}$$

since $2c>b+a\Rightarrow 1-\frac{b-c}{c-a}>0$. ■

5.2 Application: Repeated Cournot duopoly

In this section we apply previous results to a textbook example, the repeated Cournot duopoly with a linear demand function $P(Q)=\alpha-\beta Q$, where $Q=q_1+q_2$ is total quantity and symmetric constant marginal costs are denoted by C. While this is a continuous strategy stage game, some relevant strategic aspects are captured by a substructure similar to a Prisoner's Dilemma.[12] Hence we consider a reduced stage game assuming that duopolists restrict their attention to

- choosing Cournot-quantities $q^d=\frac{\alpha-C}{3\beta}$ (defect) or
- choosing symmetric joint monopoly quantities $q^c=\frac{\alpha-C}{4\beta}$ (collude).

According to our earlier notation, we denote the unique Nash equilibrium of this PD stage game – the Cournot-equilibrium – by $\omega=\left(q^d,q^d\right)$. Calculating the reaction functions and related profits yields payoff parameters $\lambda=(a,b,c,d)=(54X,81X,72X,64X)$ with $X=\frac{(\alpha-C)^2}{576\beta}$. Normalizing demand parameters so that $X=1$ we obtain the follow-

[12] A relevant difference is the possibility to punish more severely in the continuous strategy repeated game. This enlarges the feasable range of differentiating continuation payoffs which – as we pointed out in the previous section – can mitigate the problem.

ing bi-matrix form for the reduced Cournot PD-stage game

	q^c	q^d
q^c	72, 72	54, 81
q^d	81, 54	64, 64

.

The theorem shows that for $\delta \in (\underline{\delta}, \delta^*) = \left(\frac{9}{17}, \frac{19}{27}\right) \approx (.53, .7)$ there exists no collusive equilibrium of the discounted repeated Cournot duopoly PD-game that is not strictly risk dominated; and that for $\delta^* \leq \delta < 1$ there always exist some collusive equilibria that are strictly risk dominated by the Cournot-Nash equilibrium.

References

[1] Abreu, D. (1988) "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica* 56, 383-396

[2] Carlsson, H. and E. van Damme (1993), "Global Games and Equilibrium Selection," *Econometrica* 61, 989-1018.

[3] Friedman, J. (1971), "A Noncooperative Equilibrium for Supergames," *Review of Economic Studies* 38, 1-12.

[4] Harsanyi, J., and R. Selten (1988), *A General Theory of Equilibrium Selection in Games*, Boston MA: M.I.T. Press.

[5] Kandori, M.; Mailath, G. and R. Rob (1993), "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica* 61, 29-56

[6] Young, P. (1993), "The Evolution of Conventions," *Econometrica* 61, 57-84.

[7] van Damme, E. (1989), "Renegotiation-proof Equilibria in Repeated Prisoner's Dilemma, *Journal of Economic Theory* 47, 206-217.

[8] van Huyck, J., R. Battalio and R. Beil (1990), "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure", *American Economic Review 80,* 234-248

Appendix: Proof of Lemma 2 Space is too restricted to reformulate all necessary ingredients for the HS-definition of risk dominance in general games. Therefore, the following proof requires some familiarity with the HS theory for equilibrium selection. This proof rests on the proof for the so called 'Nash-product theorem' in HS (pp 214-), being appropriately adapted to the repeated PD game which is not an unanimity game. First, note that $u_i(\varphi) > 0$ and $v_i(\varphi) > 0$ imply that Γ_φ is a 'formation' of $\Gamma(s)$ in the sense of Harsanyi and Selten (HS pp 198-) since for no player i there exists a local best reply outside Γ_φ on φ_{-i} or ω_{-i}. Next, compute the so called 'bicentric prior strategies' (definition see HS) by defining expected payoffs of responding to the joint mixture $z\varphi_{-i} + (1-z)\,\omega_{-i}$ by either of the pure strategies φ_i or ω_i:

$$\begin{aligned}
\tilde{c}_i(z) \quad &: \; = U\left(\varphi_i, z\varphi_{-i} + (1-z)\,\omega_{-i}\right) = \frac{zc}{1-\delta} + (1-z)\left(a + \delta V_{i\varphi}\right). \\
\tilde{d}_i(z) \quad &: \; = U\left(\omega_i, z\varphi_{-i} + (1-z)\,\omega_{-i}\right) = z\left(b + \delta V_{i\omega}\right) + (1-z)\,\frac{d}{1-\delta}
\end{aligned}$$

Now compare these expected payoffs:

$$\begin{aligned}
\tilde{c}_i(z) \;&\geq\; \tilde{d}_i(z) \Leftrightarrow \\
z\left(\frac{c+d}{1-\delta} - a - b - \delta\left(V_{i\varphi} + V_{i\omega}\right)\right) \;&\geq\; \frac{d}{1-\delta} - a - \delta V_{i\varphi} \Leftrightarrow \\
z\left(u_i(\varphi) + v_i(\varphi)\right) \;&\geq\; v_i(\varphi) \Leftrightarrow \\
z \;&\geq\; \frac{v_i(\varphi)}{u_i(\varphi) + v_i(\varphi)}.
\end{aligned}$$

Player i's bicentric prior probabilities are given by the lengths of the subintervals $[z, 1]$ and $[0, z]$:

$$\begin{aligned}
p_i(\varphi_i) &= \frac{u_i(\varphi)}{u_i(\varphi) + v_i(\varphi)} \text{ and} \\
p_i(\omega_i) &= \frac{v_i(\varphi)}{u_i(\varphi) + v_i(\varphi)}.
\end{aligned}$$

To answer at p_{-i} by playing φ_i or ω_i yields player i the payoffs

$$\begin{aligned}
U_i(\varphi_i, p_{-i}) &= \frac{u_i(\varphi)\, u_{-i}(\varphi)}{u_{-i}(\varphi) + v_{-i}(\varphi)} \text{ and} \\
U_i(\omega_i, p_{-i}) &= \frac{v_i(\varphi)\, v_{-i}(\varphi)}{u_{-i}(\varphi) + v_{-i}(\varphi)}.
\end{aligned}$$

Comparing these payoffs yields the assertion of the lemma. ■